NOTE TECHNIQUE

AGROBASE : un système de gestion de données expérimentales

Daniel WALLACH & Jean-Pierre RELIER (*)

I.N.R.A., Laboratoire de Biométrie

() Station d'Agronomie, chemin de Borde Rouge, Centre de Recherches de Toulouse, B.P. 27, Auzeville, F 31326 Castanet Tolosan Cedex*

RÉSUMÉ

On présente les caractéristiques d'AGROBASE, un système de gestion de données expérimentales agronomiques. Ce logiciel, d'un apprentissage rapide, peut être utilisé directement sur une large gamme d'expériences. Autour d'un noyau, constitué d'un logiciel commercialisé largement répandu, on a ajouté une couche de programmes libérant l'utilisateur des contraintes syntaxiques du noyau, et enchaînant automatiquement les commandes élémentaires nécessaires à la réalisation d'une tâche donnée. Aux sorties standards fournies par AGROBASE, l'utilisateur peut adjoindre des programmes satisfaisant ses besoins particuliers. L'utilisation de ce type de logiciel peut améliorer la valorisation du travail expérimental.

Mots clés additionnels : *Base de données, informatique.*

SUMMARY

AGROBASE : a data base management system for experimental data.

A data base management system for data from agronomic experiments is presented. This software, named AGROBASE, is very easily learned, and is applicable to a wide range of experiments. Agrobases uses as a starting point a commercial data base management program, to which is added a layer of procedures which combine the elementary commands of the commercial program so that complex tasks can be accomplished by simple choice from a menu. AGROBASE contains various standard output options, to which can be added outputs specifically adapted to each user's requirements. The use of this type of software can improve the exploitation of experimental results.

Additional key words : *Data base, data processing.*

I. INTRODUCTION

L'expérimentation agronomique représente souvent un investissement important, à valoriser au mieux à partir des données recueillies.

On valorise des informations en les traitant de façons variées, ou en les stockant à des fins de statistiques pluriannuelles, ou encore en les partageant avec d'autres utilisateurs. Ces fonctions peuvent être remplies par une base de données informatisée. Encore faut-il disposer d'un logiciel pour la constituer, la gérer et l'exploiter.

Nous présentons un tel logiciel, AGROBASE, développé et utilisé à l'I.N.R.A. de Toulouse. Ce logiciel

est spécifiquement adapté aux données issues d'expériences agronomiques, mais non pas à une expérience particulière. Il peut donc être utilisé tel quel dans un grand nombre de situations d'expérimentation. Dans d'autres cas, où l'utilisation d'AGROBASE dans son intégralité ne convient pas, les choix qui y sont faits, et qui sont décrits ici, peuvent néanmoins aider à guider le développement d'un autre système de gestion de données.

II. SUPPORT INFORMATIQUE

Les choix initiaux et fondamentaux étaient :

— de développer le logiciel sur microordinateur,

pour profiter de la facilité d'accès à ces machines pour des expérimentateurs,

— de le fonder sur un logiciel commercialisé de gestion de base de données, pour disposer d'une programmation efficace des fonctions élémentaires de gestion. Le logiciel commercial choisi est KMAN (produit de Micro Data Base Systems), qui est parmi les plus performants logiciels de ce type (Anonyme, 1984). AGROBASE est donc un « superlogiciel », qui adapte KMAN aux besoins spécifiques de la gestion de données agronomiques. En particulier, pour développer AGROBASE à partir de KMAN, il a fallu décider comment les données seraient réparties et organisées dans les tables de données. Ensuite, il a fallu créer, à partir des fonctions élémentaires de KMAN, des programmes permettant d'une part de dégager le plus possible l'utilisateur des contraintes syntaxiques de KMAN, et d'autre part d'enchaîner automatiquement les commandes KMAN nécessaires à la réalisation d'une tâche donnée.

Pour installer et utiliser AGROBASE, il faut donc posséder le logiciel KMAN, et un microordinateur qui supporte ce logiciel (systèmes d'exploitation MS-DOS, PC-DOS ou CP/M-86), avec disque dur.

Les choix d'ordinateur et de logiciel ne sont pas définitifs. Il n'y a aucun problème pour transporter AGROBASE et la base de données qu'il gère d'un ordinateur à un autre. D'autre part, AGROBASE contient des procédures pour sortir toutes les données de la base sous forme de fichiers de format standard, ce qui permet de les transférer éventuellement à un autre système de gestion de base de données.

III. CARACTÉRISTIQUES D'AGROBASE

Le logiciel KMAN, et donc AGROBASE, crée et gère une base de données de type relationnel. Avec ce type de structure, les données sont stockées dans des tables rectangulaires, et chaque table contient des

informations se rapportant à un aspect particulier de l'expérience. Les tableaux 1 et 2 montrent, comme exemple, quelques enregistrements de 2 des tables d'AGROBASE.

Dans la présentation d'AGROBASE ci-dessous, on n'insiste pas sur les fonctions de ce logiciel, qui sont plus ou moins classiques (saisie de données, recherche de données d'après plusieurs critères, échanges de données avec des fichiers de format standard, ...). On présente plutôt les caractéristiques d'AGROBASE par rapport à 4 objectifs, qui sont particulièrement importants dans le contexte d'expérimentation agronomique.

A. AGROBASE accepte une grande gamme de données

nécessaire pour ne pas être limité à une expérience spécifique ou à une utilisation particulière des données.

Ceci est accompli de 2 façons. D'abord, il définit 39 tables de données qui permettent de stocker des informations concernant les opérations de gestion technique, les mesures sur le sol et sur la plante, l'organisation et l'environnement de l'expérience, le climat, et les codes utilisés (tabl. 3).

Deuxièmement, l'utilisation étendue de codes permet de stocker une grande variété d'informations dans une table donnée, en lui conservant une structure simple. Par exemple, dans la table « mesures sur le sol » le type de mesure est indiqué par un code. Ainsi, dans le tableau 1, le 1^{er} enregistrement contient le résultat d'une mesure de l'humidité du sol, le 2^e, le pourcentage de matière organique, et le 3^e, la notation d'émiettement.

Il faut souligner que la structure relationnelle de la base de données permet de regrouper facilement des informations de différentes tables. Supposons, par exemple, que l'on veuille visualiser ensemble les mesures sur le sol et les opérations de travail du sol pour la

TABLEAU 1

La table de données « mesures sur le sol ».

The « soil measurements » data table.

Expérience	Nom parcelle	Date	Code mesure	Profondeur		mesure	
				haut	bas	notation	valeur
Rotations-intrants	A01	31/08/86	humi-sol	0	10		12,2
Rotations-intrants	A01	24/09/86	mat-org	0	20		2,6
Rotations-intrants	A01	22/02/87	emiette	0	0	fort	

TABLEAU 2

La table de données « travail du sol ».

The « tillage » data table.

Expérience	Nom parcelle	Date	Code travail	Code outil	Profondeur (cm)
Rotations-intrants	A01	31/08/86	superficiel	herse	12
Rotations-intrants	A01	24/09/86	profond	chisel	30
Rotations-intrants	A01	22/02/87	profond	cover-crop	20

TABLEAU 3
Liste des tables de données.
List of data tables.

Mesures et gestions
Semis
Détails sur les semis
Récolte
Détails sur les récoltes
Travail du sol
Détails sur les travaux du sol
Fertilisation
Détails sur les fertilisations
Irrigation
Détails sur les irrigations
Traitement phytosanitaire
Détails sur les traitements phytosanitaires
Mesures sur la plante
Mesures sur le sol
Stades phénologiques
Passages du tracteur
Entités
Stations météorologiques
Détails sur les stations météorologiques
Commentaires sur les stations météorologiques
Sites
Détails sur les sites
Champs
Caractéristiques des champs
Expériences
Détails sur les expériences
Commentaires sur les expériences
Définitions des facteurs
Définitions des niveaux des facteurs
Parcelles
Détails sur les parcelles
Commentaires sur les parcelles
Niveaux des facteurs liés aux parcelles
Données météorologiques journalières
Données météorologiques journalières
Codes
Liste des attributs codés
Codes primaires
Codes secondaires
Textes sur les codes primaires
Valeurs numériques liées aux codes primaires
Valeurs codées liées aux codes primaires

parcelle A01. Il suffit pour cela de rechercher les enregistrements avec ce nom de parcelle dans les 2 tables de données.

B. AGROBASE est d'un apprentissage rapide et facile

ce qui est nécessaire pour rendre les données très accessibles à des non informaticiens.

Pour le rendre convivial, les techniques classiques ont été employées. D'abord, le logiciel est piloté complètement à partir de menus. Ainsi, chaque choix d'activité sur la base se fait en répondant, avec un numéro, aux possibilités présentées. D'autre part, au cours de la saisie des informations, des grilles de saisie sont affichées, que l'utilisateur complète. L'initiation du nouvel utilisateur à AGROBASE est facilitée par un livret de travaux pratiques inclus dans la brochure.

C. AGROBASE limite les risques d'introduire des erreurs dans la base de données

La base de données est protégée de fausses manipulations par le fait que l'utilisateur ne travaille jamais directement avec les instructions de KMAN, mais toujours par le biais d'une procédure d'AGROBASE. Ainsi il est impossible, par exemple, d'effacer par erreur une table de données, ou d'ajouter des données en oubliant de mettre à jour l'index de la table.

Pour assurer la cohérence des données stockées, AGROBASE vérifie, pour chaque nouvel enregistrement, que les codes utilisés ont été définis préalablement dans des tables prévues à cet effet, et que les dates sont acceptables (entre le début et la fin de l'expérience pour les opérations de gestion technique, entre les dates de semis et de récolte pour les mesures sur la plante). Par contre, à cause de la difficulté à définir des fourchettes partout valables, il n'y a pas de contrôle des valeurs numériques.

Si les vérifications se faisaient au moment de la saisie, des attentes pénibles en résulteraient. Pour éviter cette situation AGROBASE possède, pour chaque table de la base de données, une table « tampon » avec la même structure. L'ajout de nouvelles données se fait dans la table « tampon », sans contrôle, et donc sans problème de temps de réponse. Quand la saisie est terminée, l'utilisateur appelle la procédure qui vérifie les données et ne transfère que celles sans erreurs. (Une autre procédure permet de corriger par la suite les données erronées.)

Cette utilisation de tables « tampon » facilite d'autre part l'importation de données à partir d'un fichier classique. Il suffit de copier ces données dans la table « tampon » appropriée, puis de les traiter exactement comme des données entrées de façon interactive. Cette procédure peut être appliquée à des données saisies à l'aide d'un ordinateur portable.

D. AGROBASE est fait pour évoluer facilement

aussi bien par rapport aux données acceptées que par rapport à leur utilisation, ce qui est essentiel pour un outil de recherche.

Les données à stocker peuvent évoluer de 3 façons. S'il s'agit de nouvelles mesures sur la plante ou sur le sol, il suffit d'ajouter les nouveaux codes correspondants dans la table des codes. Deuxièmement, on peut décider d'ajouter une nouvelle table de données (comptages d'insectes, par exemple). Cela demande de la programmation, mais la structure modulaire d'AGROBASE est faite pour faciliter ce travail. Enfin, on peut être amené à modifier la structure d'une table de données existante. Un tel changement pose toujours problème, et un des objectifs lors du choix des structures des tables de données est de limiter les risques et coûts de changements de structure. Il y a des recommandations classiques pour atteindre cet objectif (MARTIN, 1977), et elles ont été adoptées pour toutes les tables d'AGROBASE.

Par rapport aux sorties, des procédures d'AGROBASE permettent d'une part d'extraire des données d'une seule table pour comparer des parcelles, et d'autre part de sortir des données de plusieurs tables qui se rapportent à la même parcelle. On peut prévoir

que chaque utilisateur aura, en plus, ses propres besoins en termes de sorties. La structure d'AGROBASE est prévue pour accueillir ces nouvelles procédures de sorties.

IV. CONCLUSIONS

Même sans compter le temps de conception et de développement d'AGROBASE, l'utilisation d'un tel logiciel sur une expérimentation nécessite un certain investissement intellectuel et matériel (temps, argent). Il existe donc un seuil en deçà duquel il est préférable de traiter les données par des procédures spécifiques, éventuellement, mais pas forcément, informatiques.

Par contre, la pérennité des dispositifs, la multiplicité des équipes travaillant sur les mêmes données, la complexité des traitements numériques, sont autant de conditions justifiant le recours à un système de gestion de base de données.

Ce mode de travail pourrait se développer parmi les chercheurs dans les années à venir. AGROBASE peut être vu comme le prototype (opérationnel) d'une nouvelle génération d'outils pour les agronomes.

Reçu le 27 janvier 1987.

Accepté le 2 juin 1987.

RÉFÉRENCES BIBLIOGRAPHIQUES

Anonyme, 1984. Project : Database, part 7. *PC magazine*, 3 (17), 218-278.

Martin J., 1977. *Computer Data-Base Organization*, 2^e éd., Prentice-Hall, New Jersey, 713 p.